AD-A069 090 PRINCETON UNIV NJ INFORMATION SCIENCES AND SYSTEMS LAB F/G 9/4
SEQUENTIAL DETECTION WITH MARKOV INTERRUPTED OBSERVATIONS.(U)
APR 79 M T HADIDI, S C SCHWARTZ N00014-77-C-0644
UNCLASSIFIED TR-43 NL

| OF |

AD
A069090

END
DATE
FILMED
6-79
DDC

Report Number 3

LEVEL II

SEQUENTIAL DETECTION WITH MARKOV INTERRUPTED OBSERVATIONS

M. T. Hadidi and S. C. Schwartz

INFORMATION SCIENCES AND SYSTEMS LABORATORY

Department of Electrical Engineering and Computer Science
Princeton University
Princeton, New Jersey 08544

APRIL 1979

Technical Report

Approved for public release; distribution unlimited

Prepared for

OFFICE OF NAVAL RESEARCH (Code 436)
Statistics and Probability Branch
Arlington, Virginia 22217

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 3	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) 6 SEQUENTIAL DETECTION WITH MARKOV INTERRUPTED OBSERVATIONS.		5. TYPE OF REPORT & PERIOD COVERED 9 Technical Report.
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) 10 M.T. Hadidi S.C. Schwartz		8. CONTRACT OR GRANT NUMBER(s) 15 N00014-77-C-0644
9. PERFORMING ORGANIZATION NAME AND ADDRESS Information Sciences & Systems Laboratory, Dept. of Elec. Eng. & Comp. Science, Princeton University, Princeton, NJ 08540		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR042-385
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research (Code 436), Department of the Navy, Arlington, Virginia 22217		12. REPORT DATE 11 April 1979
		13. NUMBER OF PAGES 47
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) 12 49p.		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

14 TR-43, 3

18. SUPPLEMENTARY NOTES

Also issued as T.R. #43, Information Sciences and Systems Lab., Dept. of EECS, Princeton University

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Sequential Detection
Markov Interrupted Observations
Switching Environment

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

We consider the sequential detection of a Markov sequence in a linear system with interrupted observations, i.e., systems with a switching environment. Because of the excessive computational requirements for optimum procedures, three suboptimum filters are discussed, all of which feed into a sequential likelihood-ratio detector. The results of a computer simulation are also presented.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

402 226

79 05 25 009

SEQUENTIAL DETECTION WITH MARKOV INTERRUPTED OBSERVATIONS

by

M.T. Hadidi and S.C. Schwartz
Department of Electrical Engineering and Computer Science
Princeton University
Princeton, NJ 08540

ABSTRACT

We consider the sequential detection of a Markov sequence in a linear system with interrupted observations, i.e., systems with a switching environment. Because of the excessive computational requirements for optimum procedures, three suboptimum filters are discussed, all of which feed into a sequential likelihood-ratio detector. The results of a computer simulation are also presented.

I. Introduction

It is common to assume linear models for the state and observation when formulating dynamical system problems. The resulting equations are simple to analyze and, provided suitable criteria are chosen for optimization, yield attractive solutions such as the well-known Kalman filter. Implicit, however, is the assumption that the origin of observations is known, which may not be true in practice. In this case the structure of the optimal solution may change completely. Such a practical aspect of the model and its consequences in problem analysis has been addressed recently (as in [1]-[7]). Here, we present results of a preliminary investigation in a related problem area.

The particular problem analyzed in this report is characterized by an observation sequence whose noise switches in a Markovian manner. Such a problem arises in multi-target tracking, [7], and was first treated by Ackerson and Fu who derived the Bayesian optimal estimator of the state, [5]. In this investigation, we focus on the detection part of the problem and present a sequential Bayesian optimal detector for the switching sequence, denoted by $\{\gamma_k\}$.

The distinction between the present problem and conventional detection problems should be made clear: here the sequence $\{\gamma_k\}$ changes according to a Markovian distribution and hence the true hypothesis switches from one stage to another. Therefore, techniques derived for problems with linear models - such as in [8] - are not applicable here, since they assume the complete observation sequence belongs to either H_0 or H_1. Furthermore, sequential detection procedures - see for example [9] - implicitly

make the above assumption and defer decision on H_0 or H_1 until time k+1, if the sequence of observation up to time k is not informative enough.

The report begins with a statement of the problem in Sec. II and then proceeds to derive the Bayesian optimal detector in Sec. III. The algorithm obtained has the nice property of being recursive; however it requires numerical integration of p.d.f.'s and hence is not practical. Therefore, we also derive three different suboptimal schemes in Sec. IV and give the corresponding detection algorithms. The results of a simulation are described in Sec. V and are followed by some observations and a comparison of the procedures in Sec. VI. In Sec. VII we make several conclusions and discuss possible extensions of this study.

II. Problem Statement and Notation

We are given a discrete-time linear system in which the measurement noise has a Markov dependent statistical property. It is described by the following equations:

$$x_k = \phi_{k,k-1} x_{k-1} + G_{k-1} u_{k-1} \tag{1}$$

$$z_k = H_k x_k + v_k + \gamma_k w_k \tag{2}$$

where x_k and u_{k-1} are vectors of dimension $n \times 1$ and $r \times 1$ while $\phi_{k,k-1}$ and G_{k-1} are matrices of the appropriate dimension. We assume that the initial state x_0 is normally distributed and the sequence $\{u_k\}$ is white Gaussian with zero mean. Thus:

$$x_0 \sim N(\cdot|\mu_0, V_0) \ , \tag{3}$$

$$u_k \sim N(\cdot|0, V_u(k)) \ , \ E\{u_j u_k^T\} = V_u(k)\delta_{j,k} \ . \tag{4}$$

The state vector enters the measurement equation, (2), linearly and is corrupted by v_k or $v_k + w_k$ depending on whether γ_k is 0 or 1, respectively. The vectors z_k, v_k and w_k are all of dimension $m \times 1$ and H_k is an $m \times n$ matrix. Again we assume $\{v_k\}$ and $\{w_k\}$ to be white noise sequences with the following statistics:

$$u_k \sim N(\cdot|0, V_v(k)) \ , \tag{5}$$

$$w_k \sim N(\cdot|0, V_w(k)) \ . \tag{6}$$

The sequence $\{\gamma_k\}$ is a binary Markov chain defined on the state space $\{0,1\}$ and is statistically described by an initial probability vector $(1-p_0, p_0)^T$ and a transition probability matrix

$$P = \begin{bmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{bmatrix} \tag{7}$$

In the above it is assumed that the vector x_0 and the sequences $\{u_k\}$, $\{v_k\}$ and $\{w_k\}$ as well as $\{\gamma_k\}$ are all mutually independent.

Our problem is to decide - at each stage k - on the value of γ_k minimizing the probability of error in detection. Formally,

$$
\begin{aligned}
&\underset{\hat{\gamma}_k}{\text{minimize}} && \Pr\{\hat{\gamma}_k \neq \gamma_k \mid z_0,\ldots,z_k\} \\
&\text{where} && \hat{\gamma}_k = \hat{\gamma}_k(z_0,\ldots,z_k) \\
& && z_k, x_k \text{ satisfy Eqs. (2) and (1) with the underlying statistics given by Eqs. (3)-(7).}
\end{aligned}
\tag{8}
$$

The derivation of the detector is described in the following section.

III. The Bayesian Optimal Detector

The minimum probability of error problem of Sec. II, is equivalent to a Bayesian decision problem in which we minimize the Bayes' risk for a special choice of the cost matrix (see for example [10]). Therefore, the problem under consideration reduces to testing - at each time k - the hypotheses:

$$
\begin{aligned}
H_1 &: z_k = H_k x_k + v_k \\
H_2 &: z_k = H_k x_k + v_k + w_k
\end{aligned}
\tag{9}
$$

where x_k evolves according to Eq. (1) and the true hypothesis switches from one stage to another according to the transition probability matrix P of Eq. (7).

By using Eq. (12), Ch. 2 of [10], the Bayesian optimal rule can be written as[1]:

$$
L_k(z^k) \triangleq \frac{f(z^k|\gamma_k=1)}{f(z^k|\gamma_k=0)} \underset{H_1}{\overset{H_2}{\gtrless}} \frac{\text{prior probability } H_1 \text{ is true}}{\text{prior probability } H_2 \text{ is true}} \tag{10}
$$

Consequently, the problem is basically that of evaluating the quantities appearing in Eq. (10). We begin with the densities $f(z^k|\gamma_k)$ and apply Bayes' rule to get:

$$
\begin{aligned}
f(z^k|\gamma_k) &= f(z_k|z^{k-1},\gamma_k)f(z^{k-1}|\gamma_k) \\
&= f(z_k|z^{k-1},\gamma_k)\frac{f(z^{k-1},\gamma_k)}{p(\gamma_k)}
\end{aligned}
$$

---

[1]For convenience we use the notation $z^k \triangleq \{z_1,\ldots,z_k\}$

$$= f(z_k|z^{k-1},\gamma_k) \times [f(z^{k-1}|\gamma_k,\gamma_{k-1} = 0)p(\gamma_k|\gamma_{k-1} = 0)\Pr\{\gamma_{k-1} = 0\}$$

$$+ f(z^{k-1}|\gamma_k,\gamma_{k-1} = 1)p(\gamma_k|\gamma_{k-1} = 1)\Pr\{\gamma_{k-1} = 1\}] \Big/$$

$$[p(\gamma_k|\gamma_{k-1} = 0)\Pr\{\gamma_{k-1} = 0\} + p(\gamma_k|\gamma_{k-1} = 1)\Pr\{\gamma_{k-1} = 1\}]$$

$$= f(z_k|z^{k-1},\gamma_k) \times f(z^{k-1}|\gamma_k,\gamma_{k-1} = 0)$$

$$\times \frac{\left[1 + \dfrac{f(z^{k-1}|\gamma_k,\gamma_{k-1} = 1)\ p(\gamma_k|\gamma_{k-1} = 1)\ \Pr\{\gamma_{k-1} = 1\}}{f(z^{k-1}|\gamma_k,\gamma_{k-1} = 0)\ p(\gamma_k|\gamma_{k-1} = 0)\ \Pr\{\gamma_{k-1} = 0\}}\right]}{\left[1 + \dfrac{p(\gamma_k|\gamma_{k-1} = 1)\ \Pr\{\gamma_{k-1} = 1\}}{p(\gamma_k|\gamma_{k-1} = 0)\ \Pr\{\gamma_{k-1} = 0\}}\right]}. \quad (11)$$

Noting that $f(z^{k-1}|\gamma_k,\gamma_{k-1}) = f(z^{k-1}|\gamma_{k-1})$ by the Markov property and defining $L_{k-1}(z^{k-1}) \triangleq f(z^{k-1}|\gamma_{k-1} = 1)/f(z^{k-1}|\gamma_{k-1} = 0)$, we obtain upon substitution from (11) into (10):

$$L_k(z^k) = \frac{f(z_k|z^{k-1},\gamma_k = 1)}{f(z_k|z^{k-1},\gamma_k = 0)} \times$$

$$\frac{\left[1 + L_{k-1}(z^{k-1})\dfrac{P(1|1)}{P(1|0)}\ \dfrac{\Pr\{\gamma_{k-1} = 1\}}{\Pr\{\gamma_{k-1} = 0\}}\right]\left[1 + \dfrac{P(0|1)}{P(0|0)}\ \dfrac{\Pr\{\gamma_{k-1} = 1\}}{\Pr\{\gamma_{k-1} = 0\}}\right]}{\left[1 + L_{k-1}(z^{k-1})\dfrac{P(0|1)}{P(0|0)}\ \dfrac{\Pr\{\gamma_{k-1} = 1\}}{\Pr\{\gamma_{k-1} = 0\}}\right]\left[1 + \dfrac{P(1|1)}{P(1|0)}\ \dfrac{\Pr\{\gamma_{k-1} = 1\}}{\Pr\{\gamma_{k-1} = 0\}}\right]} \quad (12)$$

Eq. (12) suggests that we can carry out our detection scheme in a sequential manner by using $L_{k-1}(z^{k-1})$ and $\Pr\{\gamma_{k-1} = i\}$, for $i = 0,1$, obtained at stage (k-1), and by computing the density of z_k conditioned on the observations z^{k-1} under each hypothesis. Another implication of Eq. (12) is that it reduces to Scharf's and Nolte's result, [8], when we set $p = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ as they assumed. The likelihood ratio test can now be expressed explicitly, by substituting from

Eq. (12) into Eq. (10) thus yielding:

$$L_k(z^k) =$$

$$\frac{f(z_k|z^{k-1},\gamma_k=1)}{f(z_k|z^{k-1},\gamma_k=0)} \frac{\left[1+L_{k-1}(z^{k-1})\frac{P(1|1)}{P(1|0)}\frac{Pr\{\gamma_{k-1}=1\}}{Pr\{\gamma_{k-1}=0\}}\right]}{\left[1+L_{k-1}(z^{k-1})\frac{P(0|1)}{P(0|0)}\frac{Pr\{\gamma_{k-1}=1\}}{Pr\{\gamma_{k-1}=0\}}\right]} \frac{1+\frac{P(0|1)}{P(0|0)}\frac{Pr\{\gamma_{k-1}=1\}}{Pr\{\gamma_{k-1}=0\}}}{1+\frac{P(1|1)}{P(1|0)}\frac{Pr\{\gamma_{k-1}=1\}}{Pr\{\gamma_{k-1}=0\}}}$$

$$\underset{H_1}{\overset{H_2}{\gtrless}} \frac{Pr\{\gamma_k=0\}}{Pr\{\gamma_k=1\}} \tag{10-a}$$

An Algorithm for the Sequential Detector

We now proceed to evaluate the quantities appearing in Eq. (10-a), starting with the R.H.S. Using the law of total probability and the Markov property, $p(\gamma_k)$ may be expressed as follows

$$p(\gamma_k) = p(\gamma_k|\gamma_{k-1}=0)Pr\{\gamma_{k-1}=0\} + p(\gamma_k|\gamma_{k-1}=1)Pr\{\gamma_{k-1}=1\} \ . \tag{13}$$

Next we consider the L.H.S. of Eq. (10-a). Both $L_{k-1}(z^{k-1})$ and $p(\gamma_{k-1})$ are assumed to be computed at stage (k-1). The density $f(z_k|z^{k-1},\gamma_k)$ can be evaluated using the law of total probability which gives:

$$f(z_k|z^{k-1},\gamma_k) = \int_{R^n} dx_k f(x_k|z^{k-1},\gamma_k) f(z_k|x_k,z^{k-1},\gamma_k)$$

$$= \int_{R^n} dx_k f(x_k|z^{k-1},\gamma_k) N(z_k|H_k x_k, V(k)) \ . \tag{14}$$

Here we made use of Eq. (2) and the term V(k) is either $V_v(k)$ or $V_v(k) + V_w(k)$ depending on whether $\gamma_k = 0$ or 1, respectively. If we then use the Markov property, Eq. (1), and apply the law of total probability, $f(x_k|z^{k-1},\gamma_k)$ can be expressed as:

$$f(x_k|z^{k-1},\gamma_k) = f(x_k|z^{k-1},\gamma_{k-1}=0)\cdot\frac{p(\gamma_k|\gamma_{k-1}=0)\Pr\{\gamma_{k-1}=0|z^{k-1}\}}{\sum_{i=0}^{1} p(\gamma_k|\gamma_{k-1}=i)\Pr\{\gamma_{k-1}=i|z^{k-1}\}}$$

$$+ f(x_k|z^{k-1},\gamma_{k-1}=1)\cdot\frac{p(\gamma_k|\gamma_{k-1}=1)\Pr\{\gamma_{k-1}=1|z^{k-1}\}}{\sum_{i=0}^{1} p(\gamma_k|\gamma_{k-1}=i)\Pr\{\gamma_{k-1}=i|z^{k-1}\}} \tag{15}$$

where

$$f(x_k|z^{k-1},\gamma_{k-1}=i) = \int_{R^n} dx_{k-1} f(x_k|x_{k-1},z^{k-1},\gamma_{k-1}=i)\cdot f(x_{k-1}|z^{k-1},\gamma_{k-1}=i)$$

$$= \int_{R^n} dx_{k-1} N(x_k|\phi_{k,k-1}x_{k-1},G_{k-1}V_u(k-1)G_{k-1}^T)$$

$$\cdot\ \frac{f(z_{k-1}|x_{k-1},\gamma_{k-1}=i)f(x_{k-1}|z^{k-2},\gamma_{k-1}=i)}{f(z_{k-1}|z^{k-2},\gamma_{k-1}=i)}\ . \tag{16}$$

Clearly, the quantities appearing in Eqs. (15) and (16) are either Gaussian p.d.f.'s or are available from stage (k-1), and thus $f(x_k|z^{k-1},\gamma_k)$ can be calculated recursively. It remains to compute $p(\gamma_k|z^k)$ which shall be needed for the next stage (k+1). Applying Bayes' rule we get,

$$p(\gamma_k|z^k) = \frac{f(z_k|z^{k-1},\gamma_k)p(\gamma_k|z^{k-1})}{\sum_{\gamma_k=0}^{1} \text{numerator}} \tag{17}$$

and the law of total probability then gives us:

$$p(\gamma_k|z^{k-1})=p(\gamma_k|\gamma_{k-1}=0)\Pr\{\gamma_{k-1}=0|z^{k-1}\}+p(\gamma_k|\gamma_{k-1}=1)\Pr\{\gamma_{k-1}=1|z^{k-1}\}\ . \tag{18}$$

The equations just derived constitute the basis for an algorithm that detects γ_k sequentially. Formally, it is given by:

Algorithm 0

step 1 Start with

$$f(x_0|z^{-1},\gamma_0) = f(x_0) = N(x_0|\mu_0,V_0)$$

$$f(z_0|z^{-1},\gamma_0) = f(z_0|\gamma_0) = N(z_0|H_0\mu_0,H_0V_0H_0^T + V_v(0) + \gamma_0V_w(0))$$

$$L_0(z^0) = \frac{f(z_0|\gamma_0=1)}{f(z_0|\gamma_0=0)}$$

$$\Pr\{\gamma_0=1\} = p_0 = \Pr\{\gamma_0=1|z^0\}$$

step 2 Assume we are at stage k. Then use Eqs. (16), (15) and (14) to calculate $f(x_k|z^{k-1},\gamma_{k-1})$, $f(x_k|z^{k-1},\gamma_k)$ and $f(z_k|z^{k-1},\gamma_k)$, respectively. Also compute $p(\gamma_k)$ using Eq. (13).

step 3 Calculate $L_k(z^k)$ as well as decide on γ_k using the test (10-a).

step 4 Determine $p(\gamma_k|z^k)$ from Eqs. (18) and (17) and store for the next stage together with the already computed values of $f(x_k|z^{k-1},\gamma_k)$, $f(z_k|z^{k-1},\gamma_k)$, $p(\gamma_k)$ and $L_k(z^k)$.

step 5 Set k = k+1 and go to step 2.

The above algorithm is, in principle, straightforward; however, its implementation is not so simple. This stems from the fact that Eqs. (15) and (14) call for the computation of p.d.f.'s $f(x_k|z^{k-1},\gamma_{k-1}=i)$, i = 0,1 and carrying out numerical integration. Such a computation is prohibitive, especially for systems of dimension greater than 1, as pointed out by Jaffer and Gupta in the context of a similar problem, [3].

One approach to alleviate this problem is to use a decomposition as Ackerson and Fu did in [5]. There, they expressed $f(x_k|z^k)$ as a weighted sum of Gaussian p.d.f.'s; each density corresponding to a particular realization of the switching sequence $\Gamma_k = (\gamma_0,\ldots,\gamma_j,\ldots,\gamma_k)$ with $\gamma_j = 0$ or 1. They then used a bank of 2^{k+1} Kalman filters to obtain the means and variances associated with each sequence and also derived expressions for the weights, $p(\Gamma_k|z^k)$. Though a similar decomposition can be used for $f(x_k|z^{k-1},\gamma_k)$, such an approach is not practical because the number of terms involved grows exponentially.

A closer inspection of the detection relations shows that the source of difficulty lies in $f(x_k|z^{k-1},\gamma_k)$ being non-Gaussian. In contrast, if it were Gaussian, then Eq. (14) would imply that $f(z_k|z^{k-1},\gamma_k)$ is also Gaussian and we need only compute the means and variances. In other words, we could then use a Kalman filter to provide us with the needed parameters. This simplification associated with Gaussian p.d.f.'s has been exploited before and we shall utilize it in the sub-optimal procedures to be discussed shortly.

Specifically, we shall first write $f(x_k|z^{k-1},\gamma_k)$ as

$$f(x_k|z^{k-1},\gamma_k) = \int_{R^n} dx_{k-1} f(x_k|x_{k-1}) f(x_{k-1}|z^{k-1},\gamma_k)\ , \text{ and} \tag{19}$$

then proceed in 2 steps:

(i) The functional form of $f(x_{k-1}|z^{k-1},\gamma_k)$ is approximated by $N(\cdot|\mu,V)$,

(ii) The values for μ,V are NOT chosen as the actual mean and variance, $E\{x_{k-1}|z^{k-1},\gamma_k\}$ and $\mathrm{Var}(x_{k-1}|z^{k-1},\gamma_k)$, but rather as the estimate for x_{k-1} given the measurement

$z^{k-1} = (z_0,\ldots,z_{k-1})$ and the corresponding variance and, hence, will depend on the estimation method used.

We shall investigate 3 filtering procedures, in the next section, namely:

(A) Decision-directed filtering,

(B) Linear least-mean-squared error filtering,

(C) Mean-squared error nonlinear filtering.

For each scheme, the filter equations will be derived, and the corresponding algorithm for sequential detection will be described. We then present the results of a Monte Carlo simulation performed using the three detectors in Sec. V.

IV. Derivation of the Filtering Equations and the Corresponding Detection Procedures

A. Decision-Directed Filtering

This filtering scheme assumes that the decision we make about γ_k at the $k\underline{th}$ stage, $\hat{\gamma}_k$, is correct. In other words, we make the assumption that

$$\begin{aligned}\Gamma_{k-1} &\triangleq (\gamma_0,\ldots,\gamma_j,\ldots,\gamma_{k-1}) \\ &\doteq (\hat{\gamma}_0,\ldots,\hat{\gamma}_j,\ldots,\hat{\gamma}_{k-1}) \\ &\triangleq \hat{\Gamma}_{k-1}\end{aligned}$$

and hence the p.d.f. $f(x_{k-1}|z^{k-1},\gamma_k)$ can be approximated as follows

$$\begin{aligned}f(x_{k-1}|z^{k-1},\gamma_k) &\doteq f(x_{k-1}|z^{k-1},\hat{\Gamma}_{k-1},\gamma_k) \\ &= f(x_{k-1}|z^{k-1},\hat{\Gamma}_{k-1}) \\ &\triangleq f^{(1)}(x_{k-1}|z^{k-1},\gamma_k)\,. \end{aligned} \tag{20}$$

Since $f(x_{k-1}|z^{k-1},\hat{\Gamma}_{k-1})$ is the p.d.f. corresponding to a particular realization of Γ_{k-1}, it is in fact a Gaussian density of the form $N(\cdot|x^{(1)}_{k-1|k-1}\,,\,V^{(1)}(k-1))$. Consequently, the familiar Kalman filter may be used to obtain the mean and variance as shown below:

$$x^{(1)}_{k|k} = \phi_{k,k-1}x^{(1)}_{k-1|k-1} + K^{(1)}(k)(z_k - H_k\phi_{k,k-1}x^{(1)}_{k-1|k-1} \tag{21}$$

$$V^{(1)}(k) = [I - K^{(1)}(k)H_k^T]V^{(1)}(k|k-1) \tag{22}$$

where

$$K^{(1)}(k) = V^{(1)}(k|k-1)H_k^T[H_kV^{(1)}(k|k-1)H_k^T + V_v(k) + \hat{\gamma}_kV_w(k)]^{-1} \tag{23}$$

$$V^{(1)}(k|k-1) = \phi_{k,k-1}V^{(1)}(k-1)\phi^T_{k,k-1}+G_{k-1}V_u(k-1)G^T_{k-1}\,,\; V^{(1)}(0) = V_0 \tag{24}$$

The resulting algorithm for the sequential detection of γ_k can now be stated as follows:

<u>Algorithm 1</u>

<u>step 1</u> Start with

$$f(x_0|z^{-1},\gamma_0) = f(x_0) = N(x_0|\mu_0,V_0)$$

$$f(z_0|z^{-1},\gamma_0) = f(z_0|\gamma_0) = N(z_0|H_0\mu_0,H_0V_0H_0^T + V_v(0) + \gamma_0 V_w(0))$$

$$L_0(z^0) = \frac{f(z_0|\gamma_0=1)}{f(z_0|\gamma_0=0)}$$

$$\Pr\{\gamma_0=1\} = p_0 = \Pr\{\gamma_0 = 1|z^0\}$$

<u>step 2</u> Assume we are at stage k. Then use Eqs. (20), (19) and (14) to compute $f^{(1)}(x_{k-1}|z^{k-1},\gamma_k)$, $f^{(1)}(x_k|z^{k-1},\gamma_k)$ and $f^{(1)}(z_k|z^{k-1},\gamma_k)$, respectively. Also, determine $p(\gamma_k)$ from Eq. (13). (By the Gaussian assumption, the conditional densities above are Gaussian and hence are completely specified by their means and variances.)

<u>step 3</u> Compute $L_k^{(1)}(z^k)$ as well as decide on γ_k using the test (10-a).

<u>step 4</u> Compute $x_{k|k}^{(1)}$ and $V^{(1)}(k)$ from Eqs. (21-24), and store for later use in the next stage.

<u>step 5</u> Set k = k+1 and go to step 2.

We note that an important difference between the above algorithm and Algorithm 0 is in step 4, where the detector output at stage k determines the estimator structure at the same stage.

B. <u>Linear Least-Mean-Squared Error Filtering</u>

Here we use least mean squares theory, together with a linear

constraint, to compute the mean and variance appearing in the Gaussian approximation of $f(x_{k-1}|z^{k-1},\gamma_k)$. Thus:

$$f(x_{k-1}|z^{k-1},\gamma_k) \doteq N(x_{k-1}|x^{(2)}_{k-1|k-1}\ ,\ V^{(2)}(k-1))$$

$$\triangleq f^{(2)}(x_{k-1}|z^{k-1},\gamma_k)), \tag{25}$$

where $x^{(2)}_{k|k}$ satisfies the relation

$$x^{(2)}_{k|k} = F_1(k)x^{(2)}_{k-1|k-1} + F_2(k)z_k \tag{26}$$

and F_1, F_2 are chosen to minimize $E_{x,\gamma}\{(x_k-x^{(2)}_{k|k})^T Q(x_k-x^{(2)}_{k|k})|z^k\}$. We will show that F_1 and F_2 are exactly those matrices appearing in the Kalman filter with the appropriate modification. To do so, we rewrite the system equations as follows:

$$x_k = \phi_{k,k-1}x_{k-1} + G_{k-1}u_{k-1} \tag{1}$$

$$z_k = H_k x_k + \eta_k \tag{2-a}$$

where $\eta_k = v_k + \gamma_k w_k$ and the underlying statistics are given by:

$$x_0 \sim N(\cdot|\mu_0, V_0) \tag{3}$$

$$u_k \sim N(\cdot|0,V_u(k)) \quad , \quad E\{u_j u_k^T\} = V_u(k)\delta_{j,k} \tag{4}$$

$$\eta_k \sim Pr\{\gamma_k=0\}\cdot N(\cdot|0,V_v(k)) + Pr\{\gamma_k=1\}\cdot N(\cdot|0,V_v(k)+V_w(k)), E\{\eta_j\eta_k^T\} = V_\eta(k)\delta_{j,k} \tag{5-a}$$

Clearly, the above system is in the framework of the well-known Kalman filter [11], and therefore has the following solution

$$x^{(2)}_{k|k} = \phi_{k,k-1}x^{(2)}_{k-1|k-1} + K^{(2)}(k)[z_k - H_k\phi_{k,k-1}x^{(2)}_{k-1|k-1}] \tag{27}$$

$$V^{(2)}(k) = [I - K^{(2)}(k)H_k]V^{(2)}(k|k-1) \tag{28}$$

where

$$K^{(2)}(k) = V^{(2)}(k|k-1)H_k^T[H_k V^{(2)}(k|k-1)H_k^T + V_v(k) + Pr\{\gamma_k=1\}V_w(k)]^{-1} \tag{29}$$

$$V^{(2)}(k|k-1) = \phi_{k,k-1}V^{(2)}(k)\phi_{k,k-1}^T + G_{k-1}V_u(k-1)G_{k-1}^T\ ,\ V^{(2)}(0) = V_0 \tag{30}$$

Observe that the structure of the linear LMSE estimator for x_k is independent of the higher-order statistics for $\{\gamma_k\}$. These statistics, however, enter our detection procedure through the expression of the likelihood ratio, as given by Eq.(10-a).

We may now describe the sequential scheme for the detection of γ_k - using the Gaussian approximation and linear LMSE filter - by the following algorithm:

Algorithm 2

step 1 Start with

$$f(x_0|z^{-1},\gamma_0) = f(x_0) = N(x_0|\mu_0,V_0)$$

$$f(z_0|z^{-1},\gamma_0) = f(z_0|\gamma_0) = N(z_0|H_0\mu_0,H_0V_0H_0^T + V_v(0) + \gamma_0V_w(0))$$

$$L_0(z^0) = \frac{f(z_0|\gamma_0=1)}{f(z_0|\gamma_0=0)}$$

step 2 Assume we are at stage k. Then use Eqs. (25), (19) and (14) to compute $f^{(2)}(x_{k-1}|z^{k-1},\gamma_k)$, $f^{(2)}(x_k|z^{k-1},\gamma_k)$ and $f^{(2)}(z_k|z^{k-1},\gamma_k)$, respectively. Also determine $p(\gamma_k)$ from Eq. (13). As before, only the means and variances need be evaluated for the Gaussian p.d.f.'s.

step 3 Compute $L_k^{(2)}(z^k)$ as well as decide on γ_k using the test (10-a).

step 4 Obtain $x_{k|k}^{(2)}$ and $V^{(2)}(k)$ from Eqs. (27)-(30) and store for later use in the next stage.

<u>step 5</u> Set $k = k+1$ and go to step 2.

In contrast to the D-D scheme, the estimator structure ie independent of the decision we make about γ_k and, at the same time, depends only on the first-order statistics of γ_k. As a consequence of the first fact, we expect the mean squared estimation error to be less for the linear LMSE scheme than for the D-D scheme. Further, the second fact suggests that, by incorporating the higher-order statistics of γ_k into our estimator we may be able to obtain even a better estimate. This is the case for the scheme to follow.

C. <u>Mean-Squared Error Nonlinear Filtering (Approximate Non-Gaussian Filtering)</u>

One may visualize the preceding scheme as one in which the parameters required in the Gaussian approximation of $f(x_k|z^{k-1},\gamma_k)$, are obtained as the solution of the following problem:

$$\min_{x^{(2)}_{k|k}} \quad E\{(x_k - x^{(2)}_{k|k})^T Q (x_k - x^{(2)}_{k|k})\}$$

$$\text{subject to} \quad x^{(2)}_{k|k} \text{ is a linear function in } z_k ,$$

$$z_k = H_k x_k + \eta_k ,$$

$$\text{prior of } x_k = N(\cdot|\phi_{k,k-1} x^{(2)}_{k-1|k-1} , V^{(2)}(k|k-1))$$

$$\text{prior of } \eta_k = \Pr\{\gamma_k = 0|z^{k-1}\} \cdot N(\cdot|0, V_v(k))$$

$$+ \Pr\{\gamma_k = 1|z^{k-1}\} \cdot N(\cdot|0, V_v(k) + V_w(k))$$

$$x_k, \eta_k \text{ are independent, given } z^{k-1} = \{z_0, \ldots, z_{k-1}\}.$$

It is logical, therefore, that one way to obtain a better estimate for x_k is to relax the restriction that $x^{(2)}_{k|k}$ be in the

class of linear functions in z^k. Hence by letting the estimate of x_k range over all possible functions of z^k we get an improved estimator which we shall denote by $x^{(3)}_{k|k}$ and the corresponding variance will be denoted by $V^{(3)}(k)$. The name Approximate Non-Gaussian filter, which we give to this estimator, originates from the fact that we make the incorrect assumption of a Gaussian prior for x_k which, therefore, results in approximate values for $E\{x_k|z^k\}$ and $Var\{x_k|z^k\}$. (As before, these parameters are then used in the approximation: $f(x_k|z^k,\gamma_{k+1}) = N(x_k|\mu(k),V(k))$.) The derivation of the filter uses a result due to Masreliez [12] and we give both this result and the derivation in the Appendix. The resulting estimator for the state is then described by the following equations:

$$x^{(3)}_{k|k} = \phi_{k,k-1}x^{(3)}_{k-1|k-1} + V^{(3)}(k|k-1)H_k^T g(z_k) \tag{31}$$

$$V^{(3)}(k) = [I - V^{(3)}(k|k-1)H_k^T G(z_k)H_k]V^{(3)}(k|k-1) \tag{32}$$

where

$$V^{(3)}(k|k-1) = \phi_{k,k-1}V^{(3)}(k-1)\phi^T_{k,k-1} + G_{k-1}V_u(k-1)G^T_{k-1}, V^{(3)}(0) = V_0 \tag{33}$$

$$g(z_k) = (1-q)V_1^{-1}(z_k-H_k\phi_{k,k-1}x^{(3)}_{k-1|k-1}) + qV_2^{-1}(z_k-H_k\phi_{k,k-1}x^{(3)}_{k-1|k-1}) \tag{34}$$

$$\begin{aligned} G(z_k) = {} & (1-q)V_1^{-1}+qV_2^{-1} \\ & -(1-q)q[(V_2^{-1}-V_1^{-1})(z_k-H_k\phi_{k,k-1}x^{(3)}_{k-1|k-1})(z_k-H_k\phi_{k,k-1}x^{(3)}_{k-1|k-1})^T \\ & (V_2^{-1}-V_1^{-1})] \end{aligned} \tag{35}$$

In the above we used the substitution:

$$q \triangleq \Pr\{\gamma_k = 1 | z^k\}$$

$$V_1 = H_k V^{(3)}(k|k-1)H_k^T + V_v(k)$$

$$V_2 = H_k V^{(3)}(k|k-1)H_k^T + V_v(k) + V_w(k) .$$

Observe that the equations specify an estimator which is nonlinear. This is a consequence of the observation noise being a Gaussian mixture rather than a pure Gaussian p.d.f. Also, notice that the filter structure now incorporates the higher-order statistics of $\{\gamma_k\}$ through q, and hence we expect it to perform better than the previous schemes. Based on the above filter, we get the following approximation for $f(x_{k-1}|z^{k-1},\gamma_k)$

$$f(x_{k-1}|z^{k-1},\gamma_k) \doteq N(\cdot|x^{(3)}_{k-1|k-1} , V^{(3)}(k-1))$$

$$\triangleq f^{(3)}(x_{k-1}|z^{k-1},\gamma_k) , \tag{36}$$

and the detection algorithm becomes:

<u>Algorithm 3</u>

<u>step 1</u> Start with

$$f(x_0|z^{-1},\gamma_0) = f(x_0) = N(x_0|\mu_0,V_0)$$

$$f(z_0|z^{-1},\gamma_0) = f(z_0|\gamma_0) = N(z_0|H_0\mu_0,H_0V_0H_0^T + V_v(0) + \gamma_0 V_w(0))$$

$$L_0(z^0) = \frac{f(z_0|\gamma_0=1)}{f(z_0|\gamma_0=0)}$$

$$\Pr\{\gamma_0=1\} = p_0 = \Pr\{\gamma_0=1|z^0\}$$

<u>step 2</u> Assume we are at stage k. Then use Eqs. (36), (19) and (14) to compute $f^{(3)}(x_{k-1}|z^{k-1},\gamma_k)$, $f^{(3)}(x_k|z^{k-1},\gamma_k)$ and $f^{(3)}(z_k|z^{k-1},\gamma_k)$, respectively. Also, determine $p(\gamma_k)$ from Eq. (13).

<u>step 3</u> Compute $L_k^{(3)}(z^k)$ as well as decide on γ_k using the test (10-a).

<u>step 4</u> Obtain q from Eqs. (18) and (17), and notice that the later reduces to

$$1 - q \triangleq \Pr\{\gamma_k=0|z^k\}$$

$$= \frac{1}{1 + \dfrac{\Pr\{\gamma_k=1|z^{k-1}\}}{\Pr\{\gamma_k=0|z^{k-1}\}} e^{-\frac{1}{2}(\text{residue})^T(V_2^{-1}-V_1^{-1})(\text{residue})}}$$

where the residue at k<u>th</u> stage $= z_k - H_k\phi_{k,k-1}x^{(3)}_{k-1|k-1}$. Next compute $x^{(3)}_{k|k}$ and $V^{(3)}(k)$ from Eqs. (31) - (35), and store for later use in the next stage.

<u>step 5</u> Set k = k+1 and go to step 2.

Algorithm 3 as well as the previous two algorithms are much easier to implement in comparison with the optimal detector described by Algorithm 0. To evaluate their performance, a computer simulation was performed; the results of which are presented in the following section.

V. Simulation and Results

A simulation study of the three suboptimal detection procedures was performed. The system model used is described by:

$$x_k = -\ .8x_{k-1} + u_{k-1}$$

$$z_k = x_k + v_k + \gamma_k w_k$$

where all the vectors are one-dimensional and have the following statistics:

$x_0 \sim N(\cdot|1.0,0.)$

u_k and v_k are $N(\cdot|0.,1.)$

$w_k \sim N(\cdot|0.,V_w)$

$\gamma_k \in \{0,1\}$ with transition probability matrix

$$P = \begin{bmatrix} 1-\alpha & \alpha \\ \alpha & 1-\alpha \end{bmatrix}$$

Here V_w and α are parameters which we varied in order to get different density functions for the measurement noise.

In order to simulate the derived algorithms, there were three major tasks to carry out. The first was concerned with generating the measurement sequence $\{z_k\}$, which reduced to that of obtaining the random variables involved. The Gaussian random variables were generated using the RANORM subroutine of the IBM-360 Subroutine Library, while the Markov chain, $\{\gamma_k\}$, was generated by making use of a random number generator together with the transition probability matrix (as described in [13]).

The next task was that of implementing the detection schemes.

Here, a Kalman Filter was used with appropriate modifications for each of the three filtering procedures.

The last issue to be resolved was that of evaluating the mean squared error in estimation and the probability of error in detection. These two quantities were obtained using Monte Carlo methods which provided, at the same time, information on the probability distribution of estimation error at different time instants. Specifically we used the following formulas:

$$\bar{e}_k = \text{the mean of estimation error at time } k$$
$$= E\{x_k - \hat{x}_{k|k}\}$$
$$\doteq \frac{1}{N} \sum_{i=1}^{N} (x_k^{(i)} - \hat{x}_{k|k}^{(i)}) \qquad (37)$$

where N denotes the number of runs used in the Monte Carlo simulation and the superscript (i) denotes the <u>ith</u> sample value of the appropriate random variable.

$$\overline{(e_k-\bar{e}_k)^2} = \text{the variance of estimation error at time } k$$
$$= E\{[(x_k-x_{k|k}) - E\{x_k-x_{k|k}\}]^2\}$$
$$\doteq \frac{1}{N} \sum_{i=1}^{N} [(x_k^{(i)} - \hat{x}_{k|k}^{(i)}) - \frac{1}{N} \sum_{j=1}^{N} (x_k^{(j)} - \hat{x}_{k|k}^{(j)})]^2 \qquad (38\text{-a})$$
$$= \frac{1}{N} \sum_{i=1}^{N} (x_k^{(i)} - \hat{x}_{k|k}^{(i)})^2 - [\frac{1}{N} \sum_{i=1}^{N} (x_k^{(i)} - \hat{x}_{k|k}^{(i)})]^2 \qquad (38\text{-b})$$

where the last equation follows by simple arithmetic manipulations and is introduced to simplify implementation. We finally have:

$$\text{error probability} = \Pr\{\hat{\gamma}_k \neq \gamma_k\}$$
$$= \frac{1}{N} \sum_{i=1}^{N} |\gamma_k^{(i)} - \hat{\gamma}_k^{(i)}| \;;\; \gamma_k, \hat{\gamma}_k = 0 \text{ or } 1 \qquad (39)$$

Remark: We used a value of $N = 3000$ as it proved to give sufficiently smooth curves without requiring excessive computer time.

The simulation study proceeded in two main directions. The first was to compare the performance of the three detection schemes for a specific system and under identical noise statistics. For this purpose two performance criteria were used; the mean-squared-error in estimation and the probability of error in detection. The second direction was to evaluate the performance of each scheme individually for various measurement noise distributions; in other words evaluating its sensitivity.

Results

Fig. 1(a) shows the mean-squared-error (MSE) in estimation for the Decision-Directed (DD), the Linear Least-Mean Squared Error (LLMSE) and the Approximate Non-Gaussian (ANG) filters, plotted against time. The variance of the noise sequence $\{w_k\}$ was chosen equal to the constant value 10 and the MC had the transition probability matrix, $P = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}$, which corresponds to a switching sequence $\{\gamma_k\}$ of i.i.d. r.v.'s. We observe that the (ANG) filter performs uniformly better than the (LLMSE) filter and the latter is, in turn, uniformly better than the (DD) filter. More specifically, at $k=21$ the three filters have, respectively, a MSE of 1.32, 1.47 and 1.57. In Fig. 1(b), we plotted the performance of the three detectors, as measured by the probability of error. Here we observe their performance to be surprisingly similar to one another, despite their differences in state estimation. Thus, the value of the error probability at $k=21$, is approximately 35% for all three schemes.

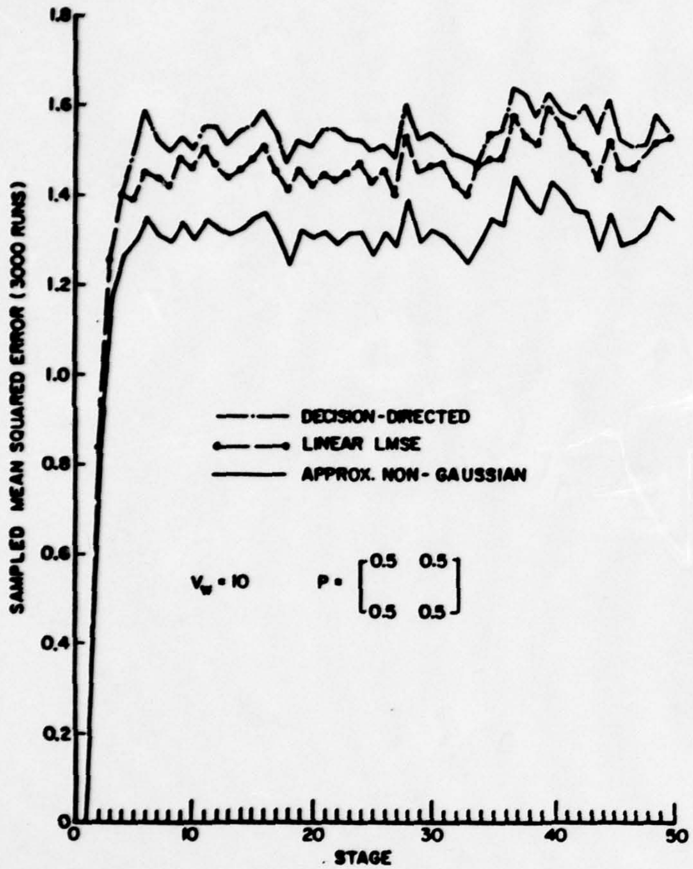

FIG. 1-a COMPARISON OF STATE ESTIMATION ERROR FOR THREE SEQUENTIAL PROCEDURES

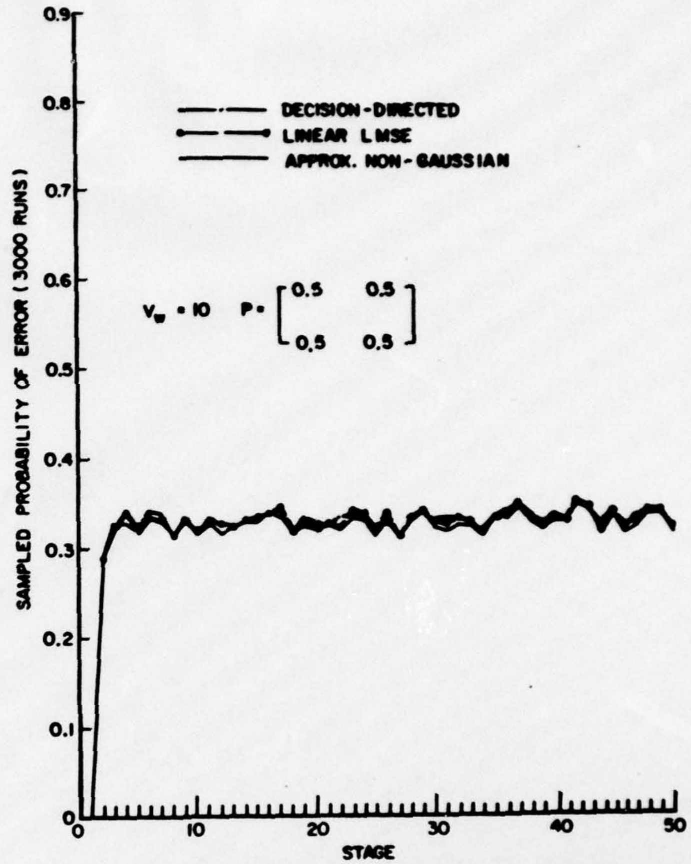

FIG. 1-b COMPARISON OF ERROR PROBABILITY FOR THREE SEQUENTIAL PROCEDURES

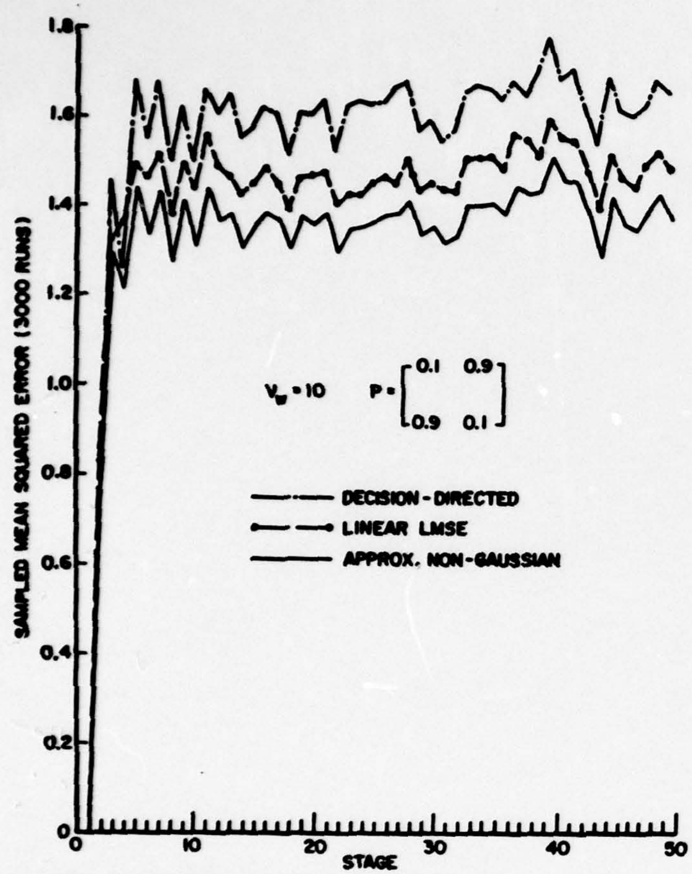

FIG. 2-a COMPARISON OF STATE ESTIMATION ERROR FOR THREE SEQUENTIAL PROCEDURES

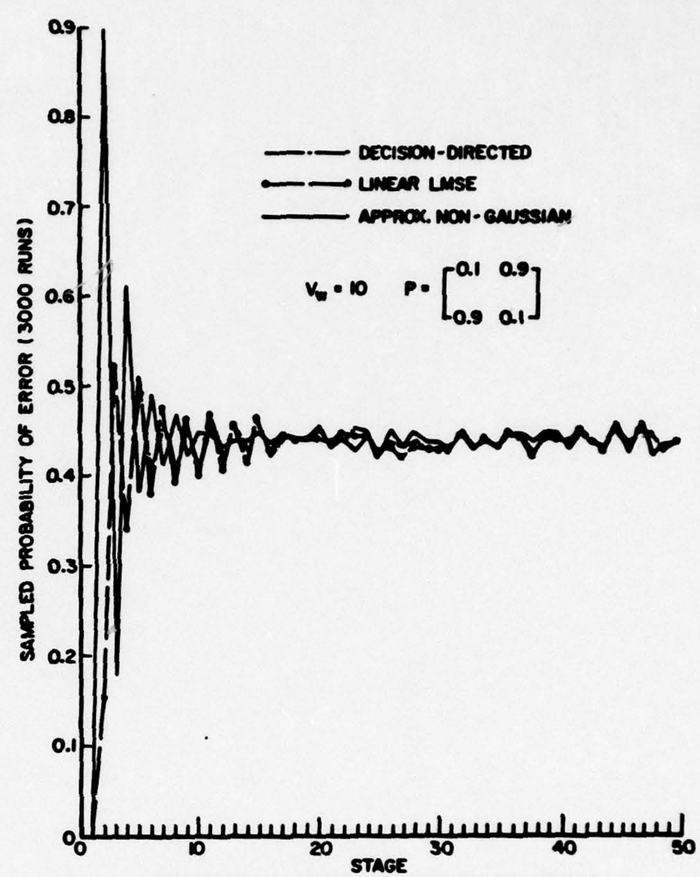

FIG. 2-b COMPARISON OF ERROR PROBABILITY FOR THREE SEQUENTIAL PROCEDURES

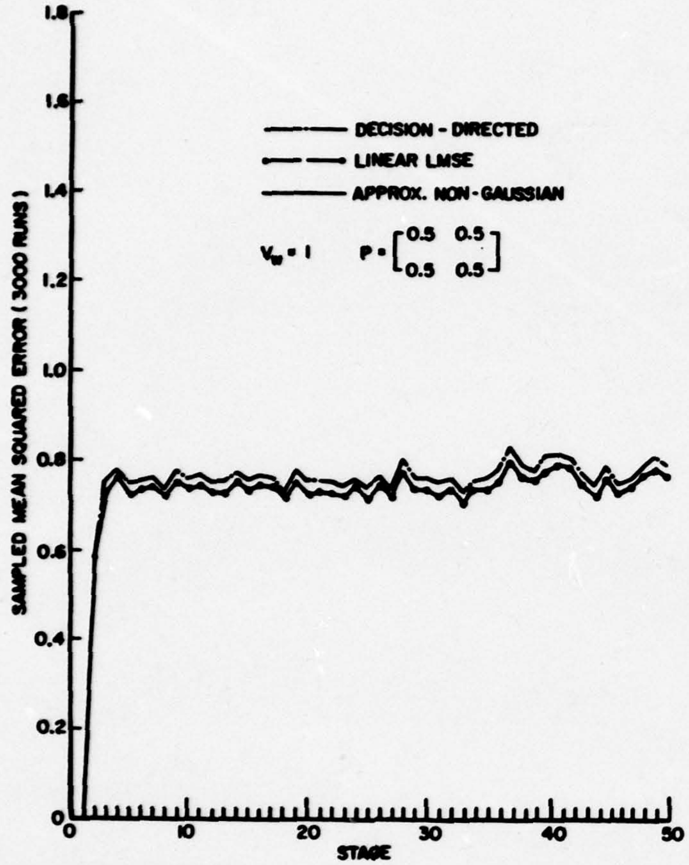

FIG. 3-a COMPARISON OF STATE ESTIMATION ERROR FOR THREE SEQUENTIAL PROCEDURES

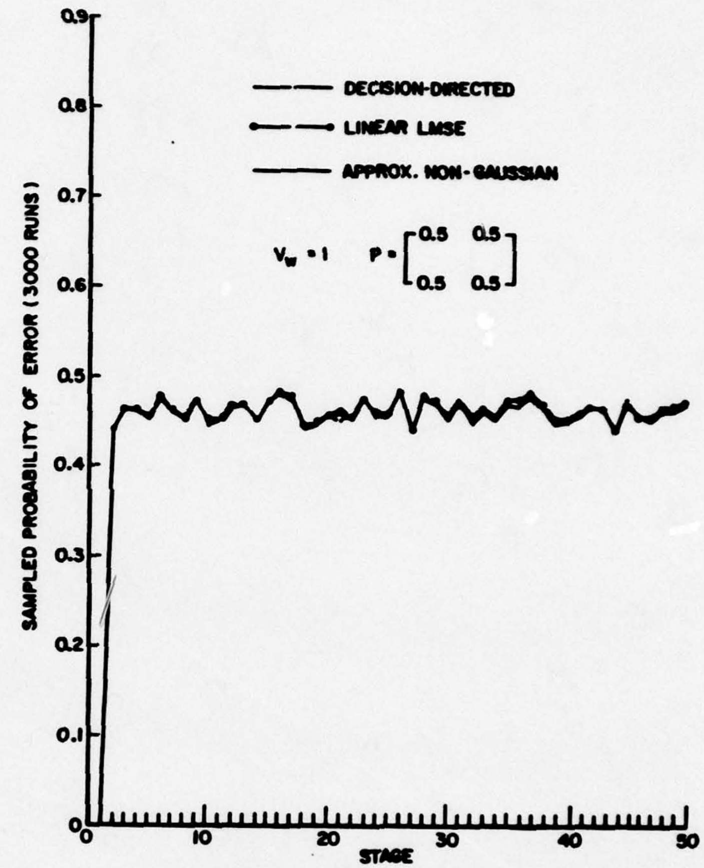

FIG. 3-b COMPARISON OF ERROR PROBABILITY FOR THREE SEQUENTIAL PROCEDURES

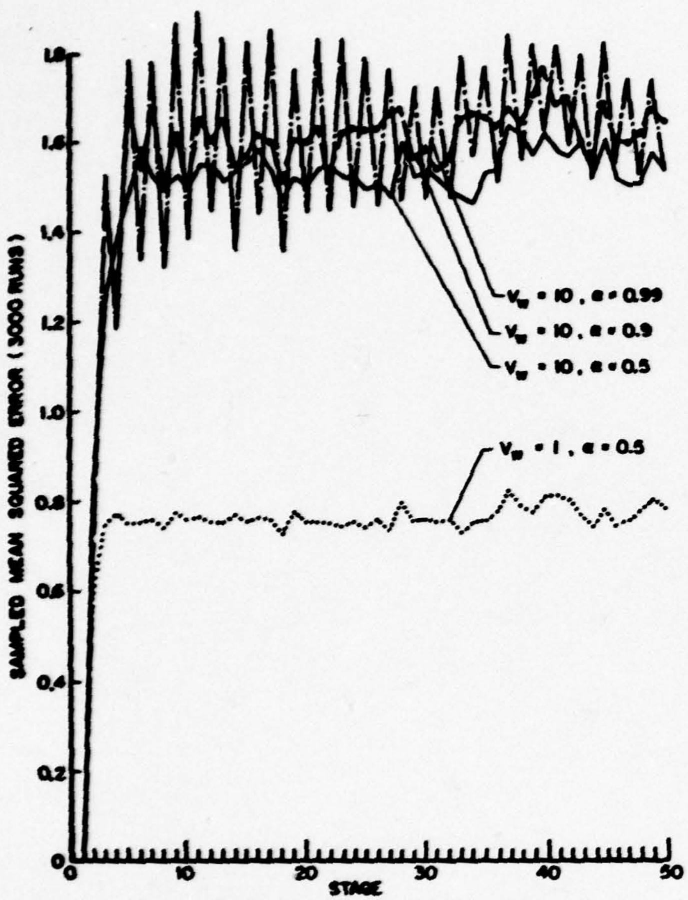

FIG. 4-a SENSITIVITY OF STATE ESTIMATION ERROR TO CHANGES IN NOISE VARIANCE, V_w, AND TRANSITION MATRIX, P. (DECISION - DIRECTED PROCEDURE)

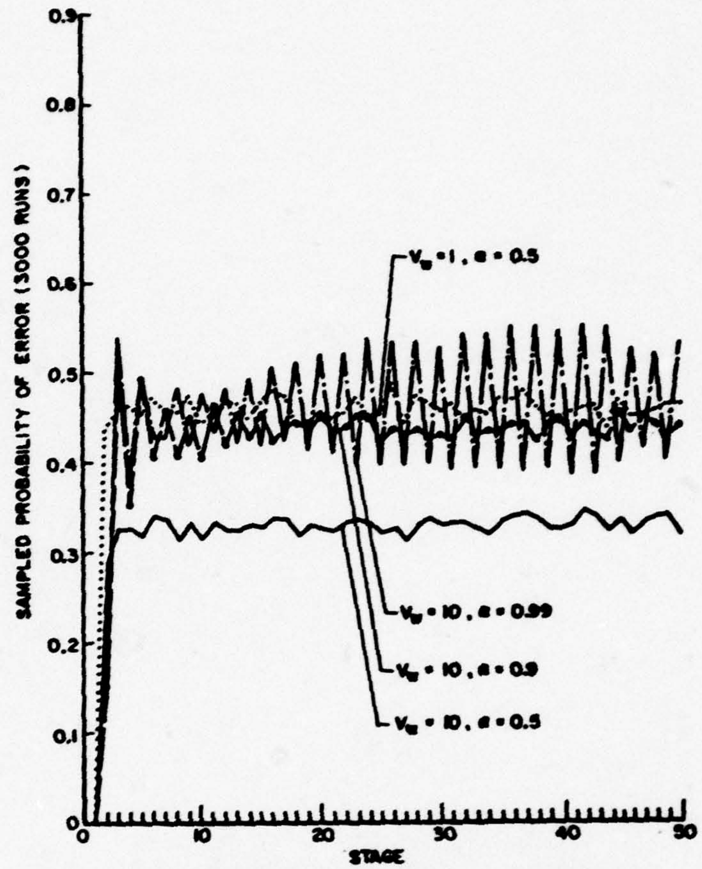

FIG. 4-b SENSITIVITY OF ERROR PROBABILITY TO CHANGES IN NOISE VARIANCE, V_w, AND TRANSITION MATRIX, P. (DECISION - DIRECTED PROCEDURE)

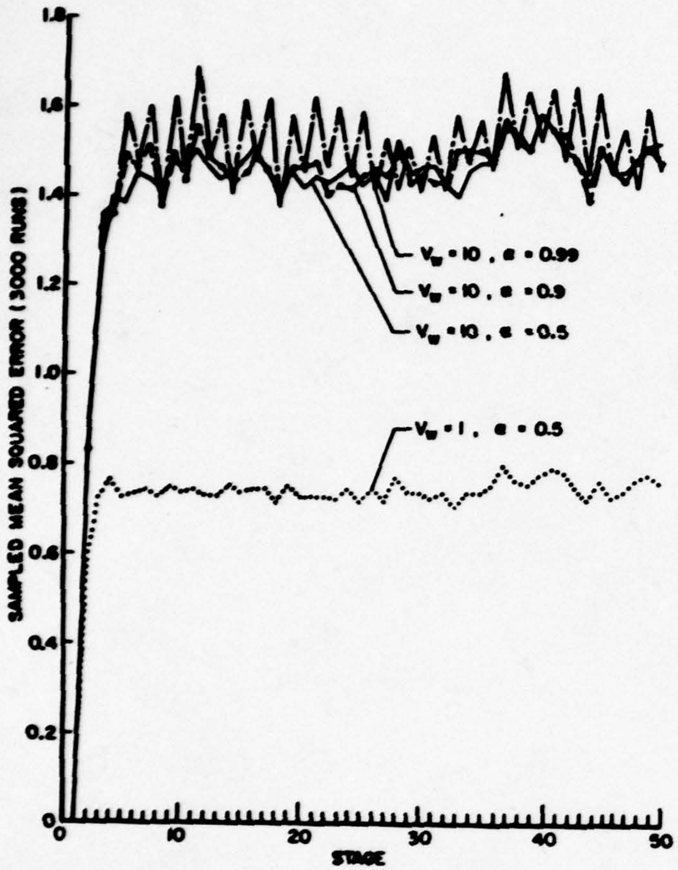

FIG. 5-a SENSITIVITY OF STATE ESTIMATION ERROR TO CHANGES IN NOISE VARIANCE, V_w, AND TRANSITION MATRIX, P. (LINEAR LMSE PROCEDURE)

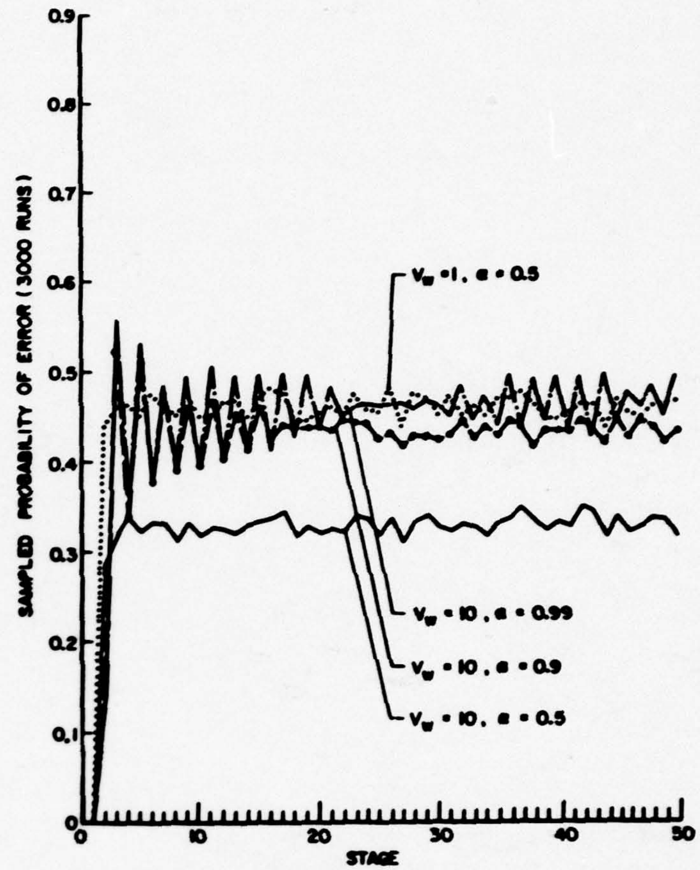

FIG. 5-b SENSITIVITY OF ERROR PROBABILITY TO CHANGES IN NOISE VARIANCE, V_w, AND TRANSITION MATRIX, P. (LINEAR LMSE PROCEDURE)

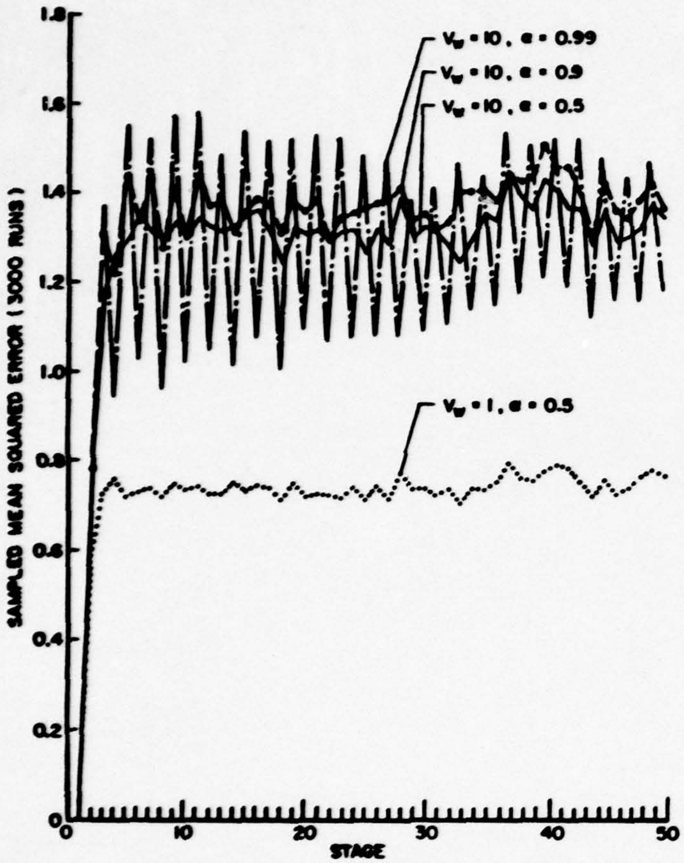

FIG. 6-a SENSITIVITY OF STATE ESTIMATION ERROR TO CHANGES IN NOISE VARIANCE, V_w, AND TRANSITION MATRIX, P. (APPROX. NON-GAUSSIAN PROCEDURE)

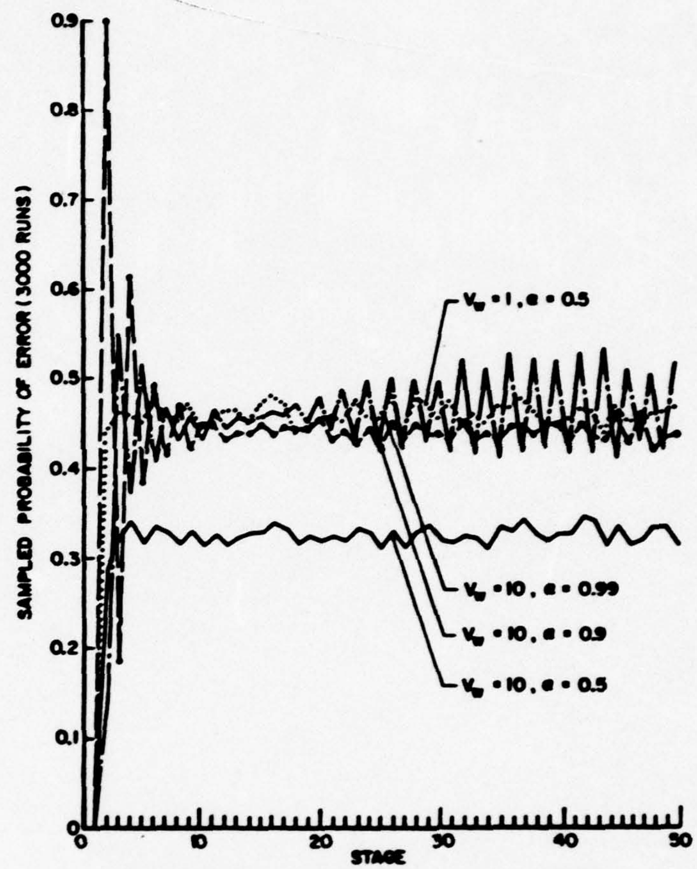

FIG. 6-b SENSITIVITY OF ERROR PROBABILITY TO CHANGES IN NOISE VARIANCE, V_w, AND TRANSITION MATRIX, P. (APPROX. NON-GAUSSIAN PROCEDURE)

Figs. 2(a,b) show the performance in estimation and detection for the three schemes with a different choice of the transition probability matrix; $P = \begin{bmatrix} .1 & .9 \\ .9 & .1 \end{bmatrix}$. Obviously our switching sequence in this case is no longer i.i.d. but rather, a highly dependent one. Nevertheless, we obtain a performance similar to the previous case with a MSE value of 1.35 for the (ANG) filter at k=21 and a probability of error of 44% at the same instant. Thus, the dependencies in $\{\gamma_k\}$ seem to bring about an increase in the MSE as well as in the probability of error. It also resulted in oscillatory transients that lasted for 10 time steps in the MSE curve and for 16 time steps in the probability of error curve.

We next examined the effect of V_w on the performance by changing V_w to 1.0 and keeping P at $\begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}$ (i.i.d. case). The results obtained are depicted in Figs. 3(a,b). We notice that the relative performance of the three schemes is similar to the case of V_w=10.0, and that the effect of reducing V_w was to reduce the MSE to .73 for the (ANG) filter and to increase the probability of error (to 45% in all three detectors). Furthermore, there is no appreciable difference in the MSE for both the (ANG) and the (LLMSE) filters, while the (DD) filter has a MSE which is higher by only .03.

We now turn to the sensitivity analysis for each scheme w.r.t. the parameters α and V_w. The results are shown in Figs. 4,5 and 6 which give both the MSE and the probability of error for the (DD), the (LLMSE) and the (ANG) procedures, respectively. Assuming a symmetric transition probability matrix, $P = \begin{bmatrix} 1-\alpha & \alpha \\ \alpha & 1-\alpha \end{bmatrix}$, we increased α from 0.5 to 0.9 and then to .99. The effect was as follows:

a) for the (DD) scheme, the MSE increased with the increase in α. Moreover, oscillatory transients were observed for $\alpha=.9$ and these oscillations persisted for $\alpha=.99$. The probability of error also increased as α increased with sustained oscillations for $\alpha=.99$.

b) for the (LLMSE) scheme, we have a similar dependency on α as in the previous scheme. However, the MSE for $\alpha=.99$ oscillates about an average value which is approximately the MSE for $\alpha=.9$, (this bias was larger for the (DD) scheme), and these oscillations have smaller amplitude. Further, the probability of error did increase as α increased, with that of $\alpha=.99$ dominating the probability of error for $\alpha=.9$ (this is different from the (DD) case).

c) for the (ANG) scheme, the general features are similar to the previous two schemes, but with the following distinctions: the MSE for $\alpha=.99$ oscillates about an average value which is less than the MSE at $\alpha=.9$, and both the oscillations in the MSE and the probability of error have amplitudes that are larger than those of the (LLMSE) detection scheme.

Finally, we investigated the effect of V_w, by reducing its value from 10.0 to 1.0, while keeping α fixed at 0.5. The resulting MSE's and probability of error's are shown in Figs. 4-6, in which we observe a common property; as V_w decreased the MSE also decreased and the probability of error in detecting γ_k, increased. We will give an explanation for this behavior as well as other observations in the next section.

VI. Discussion of Results and Comments

We shall attempt to explain some of the observations made in the previous section, starting with the relative performance of the three procedures. The results showed the (ANG) estimator outperformed the (LLMSE) estimator and the latter had less MSE than the (DD) scheme. This comes as no surprise, especially in view of our introduction to the (ANG) filter in Section IV. It was mentioned that the (ANG) filter is optimal in the MSE sense, provided the p.d.f. $f(x_k|z^{k-1})$ is Gaussian, an assumption that seems to hold as indicated by Figs. 7(a,b). On the other hand, the (DD) estimator assumes each decision we make about γ_k to be correct and determines the Kalman filter gain accordingly. Since any detector has a nonzero probability of error and in the case of the (DD) scheme there is an interaction between the estimator and the detector, then we expect incorrect decisions to propagate, resulting in a degradation of the filter performance.

The second observation is that, despite the discrepancy between the MSE of the three schemes, the probability of error is, nevertheless, practically the same. This suggests that the use of a MSE criterion for the estimator may not give the best overall detector and appears to be consistent with previously reported results [14].

Next we explain the effect of changing α. At α=.5, the switching sequence $\{\gamma_k\}$ is i.i.d. and hence the prior probability for γ_k is independent of the measurements $\{z_0,\ldots,z_{k-1}\}$, for each k. This eliminates one source of error, namely the estimation of $p(\gamma_k|z^{k-1})$. Another consequence of independence is that the state x_k and the measurement noise $\eta_k = v_k + \gamma_k w_k$ become con-

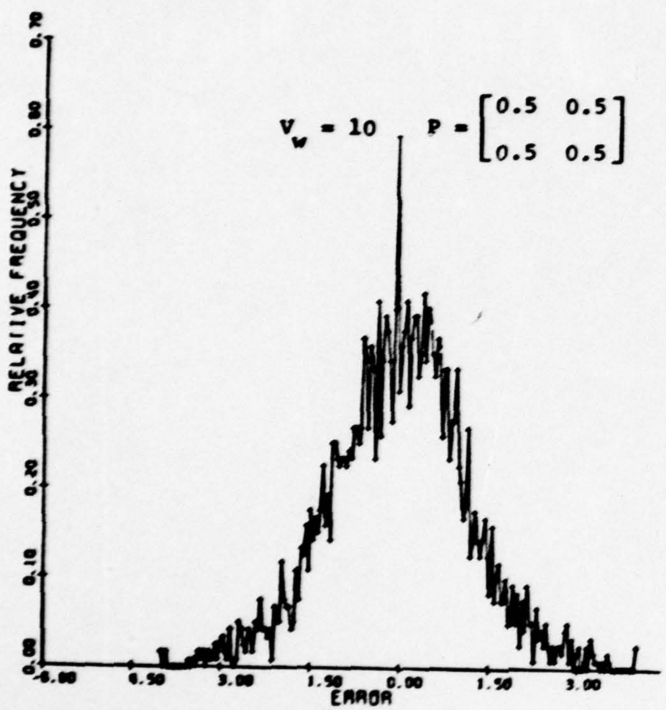

Fig. 7-a PROBABILITY DENSITY OF ERROR AT STAGE 5 FOR THE APPROXIMATE NON-GAUSSIAN PROCEDURE

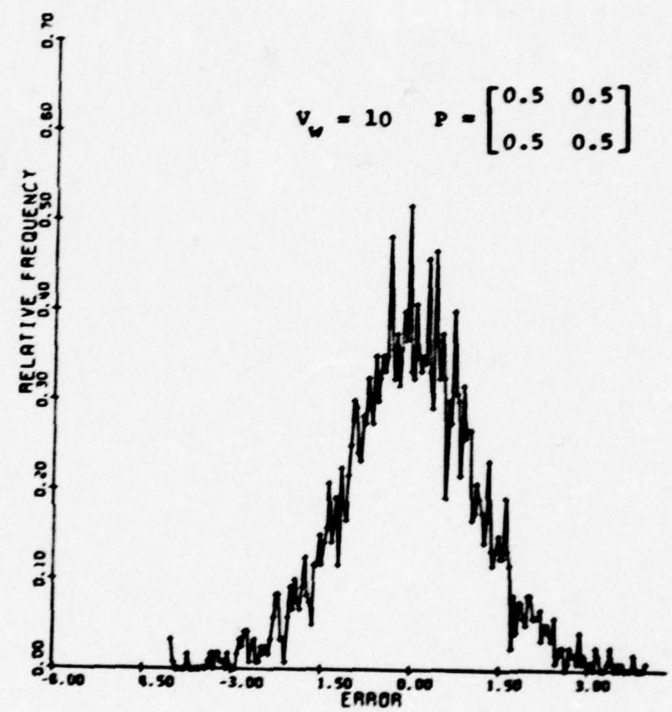

Fig. 7-b PROBABILITY DENSITY OF ERROR AT STAGE 18 FOR THE APPROXIMATE NON-GAUSSIAN PROCEDURE

ditionally independent, a requirement needed for Masreliez's theorem to hold (see the derivation in the Appendix and Ref. [12]. Due to these two factors, we expect less MSE for α=0.5 than for α=0.9 or .99. Though this seems to hold for the (DD) and the (ANG) filters, the (LLMSE) filter departs from this conclusion. For α=0.5, we also found the probability of error to be the lowest for all three schemes. This can be explained in terms of the higher accuracy in $\hat{x}_{k|k}$ and the more accurate values for $p(\gamma_k|z^{k-1})$.

Finally, we observed that by decreasing V_w, the MSE decreased while the probability of error increased. This is expected since in general $\hat{x}_{k|k}$ is a function of $\{z_0,\ldots,z_k\}$, and hence reducing the uncertainty in the measurement noise, reduces the uncertainty in $\hat{x}_{k|k}$ and therefore its variance. On the other hand, we have $z_k = H_k x_k + v_k + \gamma_k w_k$ and as the variance of w_k decreases so does the effective S/N for the detection of γ_k. Consequently, the probability of error increases.

VII. Summary and Conclusions

The problem of sequential detection in a switching environment has been investigated. Using the Markov property of the switching sequence, $\{\gamma_k\}$, and applying Bayes' rule we derived a recursive structure for the Bayesian optimal detector. The detector obtained gives a rule for deciding on the true hypothesis in a situation where the underlying hypothesis switches from one stage to another, according to a state transition probability matrix.

Since the actual implementation of the optimal detector requires numerical integration of p.d.f.'s, there is an obvious need for a more practical procedure. We undertook this task by approximating the prior p.d.f. for the state with a Gaussian density whose mean and variance are computed recursively. The three suboptimal schemes that we proposed, namely; the decision-directed procedure, the linear LMSE procedure and the approximate non-Gaussian procedure, were shown to be much simpler and easier to implement than the optimal counterpart. Moreover, the simulation study showed the decision-directed approach to be the least satisfactory while the approximate non-Gaussian was the most accurate.

An immediate application of our results is in "target tracking in a multi-target envionment." Thus, if a sensor is tracking two targets with the same state equations but different observation models, then our results provide a procedure for sequentially distinguishing the returns of one target from the other. These assumptions can easily be extended to more complicated situations. For instance, if the two targets have different state equations, thus making the model more general, we can readily modify our results by using the appropriate p.d.f.'s. Further, when there are more

than two targets, the derivation can be further modified by letting γ_k assume values in the set $\{0,1,\ldots,K-1\}$, with K the number of targets under consideration. In this case, the detection problem becomes a multiple hypothesis problem with slightly more complicated equations. Finally, since in practice we may be interested in optimizing both the decision and state estimate, we may do so by assigning costs to each aspect of the problem when initially formulating it. Such an approach was first proposed by Middleton and Esposito, [15], and is expected to yield a detector-estimator structure which is a compromise between the optimal estimator of Ackerson and Fu and our optimal detector.

We now consider the assumptions and conclusions of the above theorem and see how they apply to our problem:

1 - Masreliez's theorem gives an expression for the conditional mean estimate, $E\{x_k|z^k\}$, which is at the same time the solution to the minimum mean-squared error problem:

$$\min_{\hat{x}_k} E\{(x_k-\hat{x}_k)^T Q(x_k-\hat{x}_k)|z_k\} \quad , \quad Q \text{ is p.d.}$$

subject to

$f_x(x_k|z^{k-1}) = N(x_k|\bar{x}_k,M_k)$,

$f_\eta(\eta_k|z^{k-1})$ - is an arbitrary p.d.f. ,

x_k and η_k are conditionally independent, and

$f_z(z_k|z^{k-1})$ is twice differentiable.

This can be easily seen by expanding the quadratic term in the cost function and differentiating w.r.t. $\hat{x}_k$.

2 - By making the assumption, as we already did in our problem, that $f(x_{k-1}|z^{k-1}) = N(x_{k-1}|\hat{x}_{k-1|k-1},V(k-1))$, it follows that $f(x_k|z^{k-1}) = N(x_k|\phi_{k,k-1}\hat{x}_{k-1|k-1},V(k|k-1))$. Therefore, the requirement of the theorem that the prior of x_k be Gaussian is satisifed.

3 - If we can show that x_k and η_k are conditionally independent, and bearing in mind that z_k is actually a Gaussian mixture and hence is twice differentiable, then the remaining requirements of the theorem hold and we can apply it to our problem.

Let us now check this last step.

Appendix

Derivation of the Approximate Non-Gaussian Filter

In the following, we shall state a theorem due to Masreliez [12], which is the basis for the derivation of the approximate non-Gaussian filter. We will then compare the conditions of the theorem with the assumptions for our problem and obtain the resulting filter equations.

Masreliez Theorem

Let

$$z_k = H_k x_k + \eta_k ,$$

where

$$f_x(x_k | z^{k-1}) = N(x_k | \bar{x}_k, M_k) ,$$

$f_\eta(\eta_k | z^{k-1})$ is arbitrary p.d.f.,

x_k and η_k are conditionally independent,

and

$$f(z_k | z^{k-1}) \triangleq \int_{R^n} dx_k f_\eta(z_k - H_k x_k | z^{k-1}) f_x(x_k | z^{k-1})$$

is twice differentiable.

Then,

$$\hat{x}_k \triangleq E\{x_k | z^k\}$$

$$= \bar{x}_k + M_k H_k^T g(z_k) ,$$

where

$$g(z_k) = \frac{-\partial f_z(z_k | z^{k-1}) / \partial z_k}{f_z(z_k | z^{k-1})} \tag{A-1}$$

and

$$P_k \triangleq E\{ (x_k - \hat{x}_k)(x_k - \hat{x}_k)^T | z_k \}$$

$$= M_k - M_k H_k^T G(z_k) H_k M_k ,$$

where

$$G(z_k) = \frac{\partial g(z_k)}{\partial z_k^T} . \tag{A-2}$$

By definition $\eta_k = v_k + \gamma_k w_k$. Hence:

$$f(H_k x_k, \eta_k | z^{k-1}) = f(H_k x_k, v_k + \gamma_k w_k | z^{k-1})$$

$$= f(v_k + \gamma_k w_k | z^{k-1}) f(H_k x_k | z^{k-1}, v_k + \gamma_k w_k) \ .$$

The last term on the R.H.S. should equal $f(H_k x_k | z^{k-1})$ for conditional independence. We can write it as:

$$f(H_k x_k | z^{k-1}, v_k + \gamma_k w_k)$$

$$= \frac{f(H_k x_k, z^{k-1}, v_k + \gamma_k w_k)}{f(z^{k-1}, v_k + \gamma_k w_k)}$$

$$= \frac{f(H_k x_k, z^{k-1}, v_k, \gamma_k=0) + f(H_k x_k, z^{k-1}, v_k+w_k, \gamma_k=1)}{f(z^{k-1}, v_k, \gamma_k=0) + f(z^{k-1}, v_k+w_k, \gamma_k=1)}$$

$$= \frac{f(v_k)\Pr\{\gamma_k=0\} f(H_k x_k, z^{k-1} | \gamma_k=0) + f(v_k+w_k)\Pr\{\gamma_k=1\} f(H_k x_k, z^{k-1} | \gamma_k=1)}{f(v_k)\Pr\{\gamma_k=0\} f(z^{k-1} | \gamma_k=0) + f(v_k+w_k)\Pr\{\gamma_k=1\} f(z^{k-1} | \gamma_k=1)}$$

$$= \frac{f(v_k) f(H_k x_k, z^{k-1}) + [f(v_k+w_k) - f(v_k)]\Pr\{\gamma_k=1\} f(H_k x_k, z^{k-1} | \gamma_k=1)}{f(v_k) f(z^{k-1}) + [f(v_k+w_k) - f(v_k)]\Pr\{\gamma_k=1\} f(z^{k-1} | \gamma_k=1)}$$

$$= f(H_k x_k | z^{k-1}) \frac{\left[1 + \left(\frac{f(v_k+w_k)}{f(v_k)} - 1\right) \frac{\Pr\{\gamma_k=1\} \cdot f(H_k x_k, z^{k-1} | \gamma_k=1)}{\sum_{\gamma_k} \Pr(\gamma_k) f(H_k x_k, z^{k-1} | \gamma_k)}\right]}{\left[1 + \left(\frac{f(v_k+w_k)}{f(v_k)} - 1\right) \frac{\Pr\{\gamma_k=1\} f(z^{k-1} | \gamma_k=1)}{\sum_{\gamma_k} \Pr(\gamma_k) f(z^{k-1} | \gamma_k)}\right]}$$

It is evident from the last equation that x_k and η_k are NOT conditionally independent, in general. However, if either $f(v_k+w_k)/f(v_k)$ is equal to one or the ratio of the joint p.d.f.'s

in $H_k x_k$ and z^{k-1} appearing in the numerator and denominator are equal, then we have conditional independence. Therefore, in applying the theorem to our problem we make an error which depends on how closely the above conditions are satisfied.

The Approximate Non-Gaussian Filter for $\eta_k = v_k + \gamma_k w_k$

An inspection of Eqs. (A-1) and (A-2) shows that the filter is determined once the p.d.f. $f(z_k|z^{k-1})$ and its derivatives are evaluated. So we begin with $f(z_k|z^{k-1})$.

$$z_k = H_k x_k + \eta_k ,$$

where

$$f(x_k|z^{k-1}) = N(x_k|\phi_{k,k-1}\hat{x}_{k-1|k-1} , V(k|k-1)),$$

and

$$f(\eta_k|z^{k-1}) = f(v_k + \gamma_k w_k|z^{k-1})$$

$$= (1-p)\cdot N(\eta_k|0,V_v(k)) + p\cdot N(\eta_k|0,V_v(k)+V_w(k))$$

where we substituted p for $Pr\{\gamma_k=1|z^{k-1}\}$. It follows, therefore, that

$$f(z_k|z^{k-1}) = (1-p)\cdot N(z_k|H_k\phi_{k,k-1}\hat{x}_{k-1|k-1} , H_kV(k|k-1)H_k^T + V_v(k))$$

$$+ p\cdot N(z_k|H_k\phi_{k,k-1}\hat{x}_{k-1|k-1} , H_kV(k|k-1)H_k^T + V_v(k) + V_w(k))$$

$$= (1-p)\frac{e^{-\frac{1}{2}(z_k-\mu)^T V_1^{-1}(z_k-\mu)}}{(2\pi)^{\frac{m}{2}}|V_1|^{\frac{1}{2}}} + p\frac{e^{-\frac{1}{2}(z_k-\mu)^T V_2^{-1}(z_k-\mu)}}{(2\pi)^{\frac{m}{2}}|V_2|^{\frac{1}{2}}} \quad \text{(A-3)}$$

Here we used the substitutions

$$\mu = H_k\phi_{k,k-1}\hat{x}_{k-1|k-1}$$

$$V_1 = H_kV(k|k-1)H_k^T + V_v(k)$$

$$V_2 = H_kV(k|k-1)H_k^T + V_v(k) + V_w(k)$$

By differentiating Eq. (A-3) w.r.t. z_k we get, after some algebraic manipulations:

$$g(z_k) \triangleq -\frac{\partial f(z_k|z^{k-1})/\partial z_k}{f(z_k|z^{k-1})}$$

$$= (1-q)V_1^{-1}(z_k-\mu) + qV_2^{-1}(z_k-\mu) \tag{A-4}$$

where

$$q = \frac{\frac{p}{1-p}\frac{|V_1|^{\frac{1}{2}}}{|V_2|^{\frac{1}{2}}} e^{-\frac{1}{2}[\|z_k-\mu\|^2_{V_2^{-1}} - \|z_k-\mu\|^2_{V_1^{-1}}]}}{1+\frac{p}{1-p}\frac{|V_1|^{\frac{1}{2}}}{|V_2|^{\frac{1}{2}}} e^{-\frac{1}{2}[\|z_k-\mu\|^2_{V_2^{-1}} - \|z_k-\mu\|^2_{V_1^{-1}}]}} \tag{A-5}$$

Next we differentiate $g(z_k)$ w.r.t. z_k in order to get $G(z_k)$

$$G(z_k) \triangleq \frac{\partial g(z_k)}{\partial z_k^T}$$

$$= (1-q)V_1^{-1} + qV_2^{-1}$$

$$- (1-q)q[(V_2^{-1}-V_1^{-1})(z_k-\mu)(z_k-\mu)^T(V_2^{-1}-V_1^{-1})] \tag{A-6}$$

While $p = \Pr\{\gamma_k=1|z^{k-1}\}$ is the prior probability of γ_k given z^{k-1}, q is in fact the posterior probability of γ_k given z^k. This is seen as follows:

$$\Pr\{\gamma_k=1|z^k\} = \frac{\Pr\{\gamma_k=1|z^{k-1}\}f(z_k|z^{k-1},\gamma_k=1)}{\Pr\{\gamma_k=0|z^{k-1}\}f(z_k|z^{k-1},\gamma_k=0)+\Pr\{\gamma_k=1|z^{k-1}\}f(z_k|z^{k-1},\gamma_k=1)},$$

$$= \frac{\frac{p}{1-p}\frac{f(z_k|z^{k-1},\gamma_k=1)}{f(z_k|z^{k-1},\gamma_k=0)}}{1+\frac{p}{1-p}\frac{f(z_k|z^{k-1},\gamma_k=1)}{f(z_k|z^{k-1},\gamma_k=0)}}$$

$$= \frac{\frac{p}{1-p}\frac{|V_1|^{\frac{1}{2}}}{|V_2|^{\frac{1}{2}}} e^{-\frac{1}{2}[\|z_k-\mu\|^2_{V_2^{-1}}-\|z_k-\mu\|^2_{V_1^{-1}}]}}{1+\frac{p}{1-p}\frac{|V_1|^{\frac{1}{2}}}{|V_2|^{\frac{1}{2}}} e^{-\frac{1}{2}[\|z_k-\mu\|^2_{V_2^{-1}}-\|z_k-\mu\|^2_{V_1^{-1}}]}}$$

$$= q$$

We conclude the above analysis by noting that the filter equations just derived reduce to the familiar Kalman filter equations for the special cases, $p=0$ and $p=1$, as one would expect.

References

[1] N.E. Nahi, "Optimal recursive estimation with uncertain observations," IEEE Trans. Inform. Theory, Vol. IT-15, pp.456-462, July 1969.

[2] A.G. Jaffer and S.C. Gupta, "Recursive Bayesian estimation with uncertain observation," IEEE Trans. Inform. Theory, Vol. IT-17, pp.614-616, Sept. 1971.

[3] __________, "Optimal sequential estimation of discrete processes with Markov interrupted observations," IEEE Trans. Automat. Contr., Vol. AC-16, pp.471-475, Oct. 1971.

[4] M.T. Hadidi and S. Schwartz, "Linear recursive state estimators under uncertain observations," Tech. Rpt. #42, Information Sciences and Systems Lab., Princeton University, June 1978.

[5] G.A. Ackerson and K.S. Fu, "On state estimation in switching environments," IEEE Trans. Automat. Contr., Vol. AC-15, pp.10-17, Feb. 1970.

[6] H. Akashi and H. Kumamoto, "Random sampling approach to state estimation in switching environments," Automatica, Vol. 13, pp.429-434, July 1977.

[7] Y. Bar-Shalom, "Tracking methods in a multitarget environment," IEEE Trans. Automat. Contr., Vol. AC-23, pp.618-626, Aug. 1978.

[8] L.L. Scharf and L.W. Nolte, "Likelihood ratios for sequential hypothesis testing on Markov sequences", IEEE Trans. Inform. Theory, Vol. IT-23, pp.101-109, Jan. 1977.

[9] C.W. Helstrom, Statistical Theory of Signal Detection, Pergamon Press, London, 1968.

[10] H.L. Van Trees, Detection, Estimation and Modulation Theory, Part I, p.30, John Wiley, 1968.

[11] A.P. Sage and J.L. Melsa, Estimation Theory with Applications to Communications and Control, McGraw-Hill, 1971.

[12] C.J. Masreliez, "Approximate non-Gaussian-filtering with linear state and observation relations," IEEE Trans. Automat. Contr., Vol. AC-20, pp.107-111, Feb. 1975.

[13] S. Yakowitz, Computational Probability and Simulation, Addison-Wesley, 1977.

[14] M.H.A. Davis and E. Andreadakis, "Exact and approximate filtering in signal detection: An Example", IEEE Trans. Inform. Theory, Vol. IT-23, pp.768-773, Nov. 1977.

[15] D. Middleton and R. Esposito, "Simultaneous optimum detection and estimation of signals in noise", IEEE Trans. Inform. Theory, Vol. IT-14, pp.434-444, May 1968.

OFFICE OF NAVAL RESEARCH
STATISTICS AND PROBABILITY PROGRAM

BASIC DISTRIBUTION LIST
FOR
UNCLASSIFIED TECHNICAL REPORTS

JANUARY 1979

	Copies
Statistics and Probability Program (Code 436) Office of Naval Research Arlington, VA 22217	3
Defense Documentation Center Cameron Station Alexandria, VA 22314	12
Office of Naval Research New York Area Office 715 Broadway - 5th Floor New York, New York 10003	1
Commanding Officer Office of Naval Research Branch Office Attn: D. A.L. Powell Building 114, Section D 666 Summer Street Boston, MA 02210	1
Commanding Officer Office of Naval Research Branch Office Attn: Director for Science 536 South Clark Street Chicago, Illinois 60605	1
Commanding Officer Office of Naval Research Branch Office Attn: Dr. Richard Lau 1030 East Green Street Pasadena, CA 91101	1
Office of Naval Research San Francisco Area Office One Hallidie Plaza - Suite 601 San Francisco, CA 94102	1
Office of Naval Research Scientific Liaison Group Attn: Scientific Director American Embassy - Tokyo APO San Francisco 96503	1
Applied Mathematics Laboratory David Taylor Naval Ship Research and Development Center Attn: Mr. G. H. Gleissner Bethesda, Maryland 20084	1
Commandant of the Marine Corps (Code AX) Attn: Dr. A.L. Slafkosky Scientific Advisor Washington, DC 20380	1
Director National Security Agency Attn: Mr. Stahly and Dr. Maar (R51) Fort Meade, MD 20755	2
Navy Library National Space Technology Laboratory Attn: Navy Librarian Bay St. Louis, MS 39522	1

	Copies
U.S. Army Research Office P.O. Box 12211 Attn: Dr. J. Chandra Research Triangle Park, NC 27706	1
Naval Sea Systems Command (NSEA 03F) Attn: Miss B. S. Orleans Crystal Plaza #6 Arlington, VA 20360	1
Office of the Director Bureau of The Census Attn: Mr. H. Nisselson Federal Building 3 Washington, DC 20233	1
OASD (I&L), Pentagon Attn: Mr. Charles S. Smith Washington, DC 20301	1
ARI Field Unit-USAREUR Attn: Library c/o ODCSPER HQ USAREUR & 7th Army APO New York 09403	1
Naval Underwater Systems Center Attn: Dr. Derrill J. Bordelon Code 21 Newport, Rhode Island 02840	1
Library, Code 1424 Naval Postgraduate School Monterey, California 93940	1
Technical Information Division Naval Research Laboratory Washington, DC 20375	1

OFFICE OF NAVAL RESEARCH
STATISTICS AND PROBABILITY PROGRAM

SIGNAL PROCESSING DISTRIBUTION LIST
FOR
UNCLASSIFIED TECHNICAL REPORTS

JANUARY 1979

	Copies
Col. B. E. Clark, USMC Code 100M Office of Naval Research Arlington, VA 22217	1
Library Naval Ocean Systems Center San Diego, CA 92152	1
Professor G. S. Watson Department of Statistics Princeton University Princeton, NJ 08540	1
Professor T. W. Anderson Department of Statistics Stanford University Stanford, CA 94305	1
Professor M. R. Leadbetter Department of Statistics University of North Carolina Chapel Hill, NC 27514	1
Professor M. Rosenblatt Department of Mathematics University of California, San Diego La Jolla, CA 92093	1
Professor E. Parzen Department of Statistics Texas A&M University College Station, Texas 77840	1
Professor W. R. Schucany Department of Statistics Southern Methodist University Dallas, Texas 75275	1
Professor P.A.W. Lewis Department of Operations Research Naval Postgraduate School Monterey, CA 93940	1
Professor E. Masry Department of Applied Physics and Information Science University of California La Jolla, CA 92093	1
Professor N. J. Bershad School of Engineering University of California Irvine, California 92664	1
Professor I. Rubin School of Engineering and Applied Science University of California Los Angeles, CA 90024	1
Professor L. L. Scharf, Jr. Department of Electrical Engineering Colorado State University Fort Collins, CO 80521	1
Professor R. W. Madsen Department of Statistics University of Missouri Columbia, Missouri 65201	1

	Copies
Professor M. J. Hinich Department of Economics Virginia Polytechnic Institute and State University Blacksburg, Virginia 24061	1
Naval Coastal Systems Center Code 741 Attn: Mr. C.M. Bennett Panama City, FL 32401	1
J. S. Lee Associates, Inc. 2001 Jefferson Davis Highway Suite 802 Arlington, VA 22202	1
Naval Electronic Systems Command (NELEX 320) National Center No. 1 Arlington, Virginia 20360	1
Professor D. P. Gaver Department of Operations Research Naval Postgraduate School Monterey, California 93940	1
Professor Bernard Widrow Stanford Electronics Laboratories Stanford University Stanford, California 94305	1
Dr. M. J. Fischer Defense Communications Agency Defense Communications Engineering Center 1860 Wiehle Avenue Reston, Virginia 22090	1
Professor S. M. Ross College of Engineering University of California Berkeley, CA 94720	1

	Copies
Professor L. A. Aroian Institute of Administration and Management Union College Schenectady, New York 12308	1
Professor Grace Wahba Department of Statistics University of Wisconsin Madison, Wisconsin 53706	1
Professor Donald W. Tufts Department of Electrical Engineering University of Rhode Island Kingston, Rhode Island 02881	1
Professor S. C. Schwartz Department of Electrical Engineering and Computer Science Princeton University Princeton, New Jersey 08540	1
Professor Charles R. Baker Department of Statistics University of North Carolina Chapel Hill, NC 27514	1
Mr. David Siegel Code 210T Office of Naval Research Arlington, VA 22217	1
Professor Balram S. Rajput Department of Mathematics University of Tennessee Knoxville, Tennessee 37916	1